Dear Editors and Reviewers,

Thank you again for your careful review of our manuscript and your thoughtful comments. It is very helpful for us to see where important points in the manuscript were still unclear. We have made three main changes to the manuscript based on reviewer comments:

1. We changed Figure 1G to include genetic correlations, phenotype correlations, and heritabilities
2. We added an analysis to assess the role of local eQTL in mediation. These results are now shown in Figure 3B-C and discussed in the main text
3. We removed the eQTL results from the main text to avoid any confusion between eQTL analysis and HDMA.

Our detailed responses to each reviewer comment are below. We have included only the comments that required a response. The original R1 comment and response are included for reference and italicized. The R2 comments are in red, and the R2 responses are indented and in blue. Each new section is also annotated as to whether it is from R1 or R2 and whether it is a reviewer comment or response.

We are including the revised manuscript as well as a version with tracked changes for ease of comparison.

We think that these changes have further enhanced the clarity of the manuscript and will greatly improve reader understanding.

Gregory Carter

Reviewer #3 (Remarks to the Author):

My responses to the rebuttal are in red

*R1 review - (i) Is the permutation procedure (Fig 3B,C ) valid? It ignores relationships between* *individuals across omic levels (ie assumes all mice are exchangeable) and so might* *inflate the apparent significance. If I have understood correctly, the permutation* *procedure is performed such that it destroys any correlation between the three omic* *levels – including the existence of any QTLs or eQTLs, and this is likely too harsh a null* *hypothesis. It might be better to use multivariate generative models to simulate sets of* *genotypes, transcripts and phenotypes and evaluate performance on those rather than* *using permutation. An alternative might be to ask how unusual are the correlations* *cor(P\_C , T\_C) etc for the optimal choices of weights compared to randomly sampled* *weights using the unpermuted data, possibly augmented with a distribution fitted to the* *random correlations. In any event, that authors should justify their choice of permutation* *strategy and explain why it supports their thesis.*

*R1 response - We apologize for the lack of clarity on this critical point. The permutation test was performed by only permuting the transcriptomes. The correlations between genome and phenome were preserved in this test. Thus, we preserved the effects of any QTLs while breaking the association between eQTLs (and all other genetic effect on transcription) with QTLs. The essential goal of this procedure was to determine whether it was possible within a random transcriptome to identify a spurious transcriptomic profile that appeared to mediate the true genotype-phenotype correlations as well as our optimized transcriptomic profile in the true data. Our permutation test definitively rules this possibility out; the true transcriptome is more highly aligned with the genotype-phenotype map than expected at random. We stress that this is nontrivial, given that canonical correlation analysis (CCA)-based approaches have problems overfitting because they involved estimating the inverse of a covariance matrix and, hence, can detect such spurious components. We have added clarification of this point, which can be seen in lines 176-179 in the manuscript with tracked changes. The suggestion to consider the performance of HDMA under generative model hypotheses is an important one, and is the subject of ongoing research, but is beyond the scope of this paper. In this study, we use HDMA in a purely descriptive mode.*

*We randomly sampled the unpermuted weights 10,000 times and recalculated the correlations cor(G\_C, T\_C) and cor(T\_C, P\_C) as shown in Figure 3C. The results are shown below. We think it would add more confusion than information to add this to the manuscript figure, but shows clearly that the correlations identified by HDMA are well outside the null distribution. The original permutated values from the manuscript in Figure 3 are shown in gray, and the observed correlations are shown in red.*

R2 review - Thanks for clarifying the permutation procedure. So to summarise, the relationship

between genotype and phenotypes (ie QTL) is preserved by permutation but there no longer any relationship between transcriptome and either genotype or phenome, (but presumably the internal gene expression network correlation structure is preserved). Under this permutation procedure we would not expect to see any mediation, therefore. Whilst I understand the rationale for this approach, it is very severe. It does not really address the question of whether the mediation observed (ie the correlation of T\_C and P\_C) is unusually strong, given the presence of eQTLs; is T\_C better correlated with P\_C than might be expected given the existence of lots of eQTLs.

At the very minimum, the manuscript should be very clear as to what the permutation procedure is demonstrating and its limitations. Ideally a different approach should be adopted, although I understand that this would be a significant amount of work and quite difficult.

R2 response - We apologize that we have not been able to address this point satisfactorily. To further clarify this procedure, we have added an analysis, the results of which have been added to Figure 3B and C.

We generated two additional transcriptomic kernel matrices to consider the effects of the presence of eQTL in the system. 1) We generated a "distal kernel" by first regressing out the effect of local haplotype on each transcript. We used the residuals to generate a kernel matrix that captured only distal effects on transcription. 2) We also generated a "local kernel" by imputing gene expression from local haplotype only. This included only the local effects (local eQTL) on transcription. We ran HDMA using each kernel in turn.

The path coefficient identified using the local transcriptomic kernel was not significantly different from the null, indicating that when local effects alone are considered, the transcriptome does not mediate the effects of the genome on phenome. In contrast, the path coefficient identified using the distal transcriptomic kernel, was highly significantly different from the null distribution, and indistinguishable from that identified using the full transcriptome.

Looking at the G-T and T-P correlations, we see that the G-T correlation is only slightly lowered by using the distal kernel. It should be a little lower because we have removed some of the genetic effect (local effects) of the genome on the transcriptome. The G-T correlation for the local kernel is extremely high, which also makes sense, because we have defined the transcriptome using the genome. However, the T-P correlation when using the local kernel is very low (0.14), suggesting that animals sharing many local eQTLs are not expected to be very similar in traits.

We have added text to the manuscript and methods to explain this new analysis.

Minor Points:

*R1 review - (i) Figure 1 is very good, except Figure 1G could be improved if the upper triangle of the heatmap displayed genetic correlations and the main diagonal heritabilities. This is an interesting way to add more information to the plot. We tried it (shown below), but then realized that it might be confusing to have two different types of data in a single heat map. Because most of the cells show Pearson correlation, but the diagonal has very low values, we think this might cause more confusion, especially since the heritabilities are shown in Figure 1F.*

*R1 response - This is an interesting way to add more information to the plot. We tried it (shown below), but then realized that it might be confusing to have two different types of data in a single heat map. Because most of the cells show Pearson correlation, but the diagonal has very low values, we think this might cause more confusion, especially since the heritabilities are shown in Figure 1F.*

R2 review - The Figure shown above in the response does not correspond to what was requested – I was asking for the upper triangle to be the genetic correlations and the lower triangle the phenotypic correlations, in addition to the main diagonal being the heritabilities. The Figure seems to have just the genetic correlations.

R2 response - We apologize for misunderstanding the initial comment. We think we understand better now what is being asked, and we have revised figure 1G. We now use a tri-color heat map to show the genetic correlations, phenotype correlations, and heritabilities for the traits. This new panel is pasted below.



*R1 review - (ii) Figure 2 could be significantly improved by replacing the violin plots in Fig2A with*

*overlapping distributions as in Figure 1A,B. Fig 2B is informative but I think the use of*

*linear regression is not the best way of showing the shapes of the distributions for Local*

*and Distal are different. Clearly a straight line does not fit any of the data very well. Can*

*the authors think of another measure which quantifies the fact that strong local eQTL*

*are more likely to have small trait correlations than strong distal eQTL?*

*R1 response - We tried using overlapping distributions (below), but it became difficult to*

*compare the distributions across the tissues. We would like Figure 2A to show a*

*direct comparison of the variance explained by local and distal eQTLs across*

*tissues, so we reverted to the violin plots.*

R2 review - This is a matter of personal preference, so I don’t insist on the use of overlapping distributions, but I do find the violin plots to be so similar it is hard to draw any conclusions about the distributions. Incidentally the Figure legend now seems to be adjusted for the overlapping distribution plots, not the violin plots.

R2 response - Thank you for catching this error. We have fixed the legend.

*R1 review - (iv) Figure 4 D – Please include scatter plots of P\_C vs T\_C, G\_C vs T\_C, G\_C vs P\_C - this is surely key to understanding. Please replace violin plots with overlapping distributions as in Figure 1A,B*

*R1 response - We tried using overlapping histograms to compare these distributions, but the distributions are so similar that the figure was difficult to read (see below). Instead, we replaced the distributions with separate panels showing the observed loading distribution for each tissue compared with the null. We shaded the area of the distributions that were more extreme than the null distribution and noted the number of genes in each extreme group. This adds a little more information to help distinguish the distributions. The scatter plots took up a lot of room in Figure 3 without adding much information, so we have created a new Supplemental Figure S3 to show the details of these correlations.*

R2 review - I think the scatter plots are an excellent addition and should be in the main paper and the violin plots/overlapping distributions in the supplement.

R2 response - We have added the scatter plots to figure 4 and added the overlapping distributions as Supplemental Figure 3.

*R1 review - (ix) What were the exact criteria for calling eQTL (local and distal) and pQTL? Were different significance thresholds applied for local vs distal? The methods are vague – why does a LOD score threshold of 8 equate to a p-value of 0.05 (and what does this p-value mean – it is genome wide significance?) Surely it would be better to use an FDR-based threshold.*

*R1 response - Local eQTLs were defined as eQTLs within 4Mb of the transcription start site of the encoding gene. We have added this definition to the methods. We used a nominal p value threshold of 0.05 as a permissive and arbitrary cutoff to compare basic stats of local and distal eQTLs. The eQTLs were a marginal component of the manuscript (only presented in supplemental figures), and changing this threshold did not change the general conclusions about eQTLs that we presented. Namely, local eQTLs outnumber distal eQTLs and local eQTLs tend to be shared across tissues whereas distal eQTLs tend to be tissue specific. This study did not include protein data and we therefore do not report any pQTLs.*

R2 review - I don’t think a 5% cutoff is at all reasonable. The manuscript does not even report the threshold or the numbers of QTL in the main text – they are buried in supplemental Figure S2B-E, so I think this point still requires attention. By pQTL I meant physiological QTL, not protein QTL, sorry I should have made that clear.

R2 response - We apologize again for the lack of clarity in our response. We also misspoke slightly about the severity of the threshold. The LOD score of 8 corresponds to a permutation-based genome-wide threshold < 0.01. This range threshold is standard practice (PMID 7851788, ref 10 in Methods) and the islet data in this paper have been previously published with a less stringent threshold (PMID 29567659, ref 12).

In light of this comment, we think that we put too much emphasis on the eQTL analysis in this manuscript, and that inclusion of these results will be very confusing to many readers. The purpose of this manuscript is to describe high-dimensional mediation, which is completely independent of QTL mapping of any kind. We performed eQTL mapping only because we thought people would be curious about these basic results. However, in the primary analysis we did not do any mapping. To clarify this point in the manuscript, we have removed the description of the eQTL analysis from the main text and altered the language to better guide the reader through the reasoning of how we used gene expression and local and distal genetic contributions to gene expression.

*R1 comment - (ii) It not clear whether the CC-RIX mice were kept in the same animal facility as the DO mice or were from a different experiment. Were they on the same high-fat diet as the DO? - please clarify.*

*R1 response - To clarify the mouse populations, we have added "CC-RIX" to one of the headers in the methods to indicate that this section describes the CC-RIX mice (Line 40 in Methods with tracked changes). The CC-RIX mice were housed at The Jackson Laboratory. The DO mice were part of a previous experiment and were housed at the University of Wisconsin. DO and CC-RIX mice were maintained on different high-fat, high-sugar diets. DO mice received a HF/HS diet (44.6% kcal fat, 34% carbohydrate, and 17.3% protein) from Envigo Teklad (catalog number TD.08811). The CC-RIX mice received a custom-designed high-fat, high-sugar (HF/HS) diet (Research Diets D19070208).*

R1 review - These data should be included in the Methods.

R2 response - These data are included in the methods lines 8-12 and lines 41-51

*R1 review - (iv) TWAS analysis (methods line 250 onwards. Using just the SNP closest to the TSS for each gene might result in underestimating local genetic effects – it would have been better to have taken the most associated SNP within say 100kb. Not all cis SNPs will be associated with the expression trait, so picking one based solely on location is sub-optimal.*

*R1 response - We agree that this method would be problematic in human data. However, the mice used in this experiment have large haplotype blocks, and the markers within 100kb of any given marker will have identical or nearly identical genotype distributions across the animals.*

R2 review - I don’t agree – it is not true that nearby SNPs will always be surrogates for each other,

unless they also share the same strain distribution pattern in the founders.

R2 response - We apologize for a lack of clarity. We did not perform association mapping with SNPs. When working with DO and CC mice, it is standard to perform mapping with ancestral haplotype probabilities. LD blocks of the haplotypes degrade very slowly. As an example, we have pasted below the correlations between markers on Chr 1 at increasing intervals up to 500 kb (bars indicate standard deviation). Even at this distance, the correlation between markers (using haplotypes) is 0.9. The haplotypes closest to the TSS of the gene are an excellent estimate of the haplotypes in the neighborhood. This is one of the reasons we think we see much higher contributions of local genotype to transcript abundance in these mice than is typically seen in humans (discussed in lines 118-130).

